

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

H04N 7/14

H04M 11/06



# [12] 发明专利申请公开说明书

[21] 申请号 02127523.8

[43] 公开日 2003 年 1 月 22 日

[11] 公开号 CN 1392734A

[22] 申请日 2002. 5. 29 [21] 申请号 02127523.8

[30] 优先权

[32] 2001. 5. 29 [33] JP [31] 161031/2001

[71] 申请人 日本电气株式会社

地址 日本东京都

[72] 发明人 中村欣贵 正岛直树

[74] 专利代理机构 中国专利代理(香港)有限公司

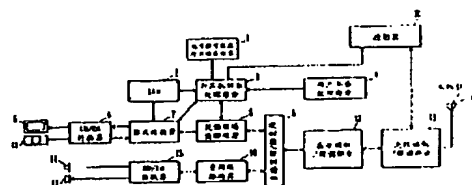
代理人 张志醒

权利要求书 2 页 说明书 9 页 附图 8 页

[54] 发明名称 电视电话设备

[57] 摘要

一种 TV 电话设备, 包括一个图象传感部件, ROM, CG 处理部分以及无线电天线。图象传感部件拍摄讲话者的真实影像。ROM 预先存储讲话者至少一段代用影像。CG 处理部分从 ROM 中提取代用影像数据, 以替换图象检测部件输出的真实影像数据, 并执行图象处理。无线电天线将 CG 处理部分输出的代用影像数据发送给语音通信的参与者。



BEST AVAILABLE COPY  
BEST AVAILABLE COPY

1. 一种电视电话设备，其特征在于，包括  
    图象传感装置（5），用于拍摄讲话者的真实影像；  
5      第一存储装置（1），用于预先存储讲话者的至少一个代用影像；  
    处理装置（3），用于从所述第一存储装置中提取代用影像数据，以代  
替所述图象传感装置输出的真实影像数据，并执行图形处理；以及  
    传送装置（11，13），用于将所述处理装置输出的代用影像数据发送到  
语音通信的参与者。
- 10      2. 根据权利要求1所述的设备，其特征在于，所述处理装置选择提取自  
所述第一存储装置的代用影像数据和所述图象传感装置输出的真实影像数据  
中的一个，并根据讲话者的指令来执行图形处理。
3. 根据权利要求1所述的设备，其特征在于，传送到语音通信参与者的  
代用影像数据包含安排在影像中的字符串数据。
- 15      4. 根据权利要求1所述的设备，其特征在于，  
    所述第一存储装置预先存储多个代用影像数据；以及  
    所述处理装置从存储在所述第一存储装置中的多个代用影像数据中选择  
要被发送的代用影像数据，并根据讲话者的指令来执行图形处理。
5. 根据权利要求1所述的设备，其特征在于，  
20      所述设备进一步包括第二存储装置（1），用于预先存储多个部分影像  
数据，作为代用影像的成分，以及  
    所述处理装置从所述第二存储装置所存储的部分影像数据中选择必要的  
影像数据，根据讲话者的指令来合成代用影像数据，并将合成的代用影像存  
储在所述第一存储装置中。
- 25      6. 根据权利要求5所述的设备，其特征在于，所述处理装置根据讲话者  
的指令来改变每个被选部分影像的颜色。
7. 根据权利要求5所述的设备，其特征在于，所述处理装置根据讲话者  
的指令来确定每个部分影像在代用影像中的位置。
8. 根据权利要求1所述的设备，其特征在于，进一步包括：

转换装置(7)，用于将提取自所述第一存储装置的代用影像数据和所述图象传感装置输出的真实影像数据转换成数字影像信号格式，以及

压缩装置(8)，用于压缩所述转换装置输出的数字影像信号，并将该信号输出到所述传输装置。

5 9. 根据权利要求8所述的设备，其特征在于，所述设备进一步包括：

接收装置(11, 13)，用于接收来自语音通信参与者的压缩影像数据；

解扩装置(8)，用于展开所述接收装置输出的压缩影像数据并输出一个数字影像信号；以及

显示装置(5)，用于显示接收到的影像；以及

10 所述转换装置转换所述解扩装置输出的数字影像信号的格式，并将该信号输出到所述显示装置。

10. 根据权利要求1所述的设备，其特征在于，所述处理装置将从所述第一存储装置中提取的代用影像数据转换成数字影像信号格式。

## 电视电话设备

5      技术领域

本发明涉及一种电话设备，该设备使用语音和视频执行远地之间的传输。

背景技术

10      充当远地之间通信工具的传统电话设备只起语音传输装置的作用，并不是以作为表情传输装置。近来，既能使用语音又能使用视频通信的TV（电视）电话装置被使用。

15      由于TV电话不仅传输语音，它还将说话者的外貌和背景作为视频显示给另一方，因此所有的呼叫者都可以知道说话者的隐私。特别的，当一个人突然被呼叫，并且他不想显示其外表和背景时，他必须选择只使用声音，而不传送任何影像的语音通信。

20      作为一种替换的方法，日本专利平一开9-331509或2000-175168中提出了一种TV电话设备，其中例如字符图像的一段替换的影像被发送，而不发送真实影像，由此使用该图像发送了讲话者的表情，同时保护了其隐私。然而，在这种TV电话设备中，由于代用影像是从个人计算机等设备中加载的，因此电话设备自身无法编辑并产生所要传送的代用影像。

发明内容

本发明的一个目的是提供一种能够使用图像来发送讲话者表情的TV电话设备，同时保护其隐私。

25      本发明的另一个目的是提供一种TV电话设备，它能够编辑和产生由电话设备自身所传输的代用影像。

为了实现上述目的，根据本发明，提供了一种TV电话设备，它包括图像传感装置，用于拍摄讲话者的真实影像，第一存储装置，用于预先存储讲话者至少一段代用影像，处理装置，用于从第一存储装置中提取代用影像数据，替

换图像传感装置所输出的真实影像数据, 并执行图象处理, 以及传输装置, 用于向语音通信中的另一方传送处理装置所输出的代用影像数据。

### 附图说明

- 5 图 1 是根据本发明实施例的便携式 TV 电话设备的方框图;  
图 2 是显示发自图 1 所示便携式 TV 电话设备的呼叫过程的流程图;  
图 3 是显示图 1 所示便携式 TV 电话设备中呼叫接收过程的流程图;  
图 4 是显示从图 1 所示便携式 TV 电话设备发送一段真实影像到另一方终端的过程的流程图;  
10 图 5 是显示从图所示 1 便携式 TV 电话设备发送一段代用影像到另一方终端的过程的流程图;  
图 6 是显示代用影像定义数据格式的视图;  
图 7 是显示代用影像生成窗口的视图;  
图 8 是显示当生成窗口上的工具条被操作时的一种状态的视图;  
15 图 9A 和 9B 是显示语音通信期间发送方和接收方 TV 电话设备的状态的视图;  
图 10 是显示语音通信期间 TV 电话设备显示屏的视图; 以及  
图 11 是显示图 1 所示便携式 TV 电话设备中的接收过程的流程图。

### 20 具体实施方式

以下将参考附图对本发明进行详细的描述。

- 图 1 示出了根据本发明实施例的便携式 TV 电话设备。参考图 1, 一个 EEPROM (电可擦可编程序只读存储器) 1 存储图形数据, 一个 RAM2 存储临时数据。  
25 一个 CG (计算机制图) 处理部分 3 处理从 ROM1 中读出的图形数据, 经 CG 处理部分 3 处理的数据被一个格式转换器 7 或一个视频编解码器 8 所使用。用户界面处理部分 4 将来自用户的指令传送到 CG 处理部分 3。例如液晶显示板等显示设备 5 对 AD/DA 转换器 6 的模拟输出影像进行显示。

AD/DA 转换器 6 将来自格式转换器 7 的数字信号转换成可被显示设备 5 使

用的模拟信号,还将来自图像传感部件 10 的模拟信号转换成可被格式转换器 7 使用的数字信号。格式转换器 7 对 AD/DA 转换器 6、视频编解码器 8 以及 CG 处理部分 3 之间的信号进行转换。视频编解码器 8 压缩或解扩视频信号。

5 多路复用/解复用部分 (MUX/DEMUX) 9 对来自视频编解码器 8 的视频信号和来自音频编解码器 16 (以后将被描述) 的声音信号进行多路复用,并将该多路复用信号传送到基带调制/解调部分 12。多路复用/解复用部分 9 还将来自基带调制/解调部分 12 的接收信号解复用成声音和视频信号。图像传感部件 10 拍摄用户图像并输出一个模拟视频信号。无线电天线 11 发送/接收无线电波。

10 基带调制/解调部分 12 对来自多路复用/解复用部分 9 的信号进行基带调制,并将信号输出到一个无线调制/解调部分 13。基带调制/解调部分 12 还对来自无线调制/解调部分 13 的信号进行基带解调,并将信号输出到多路复用/解复用部分 9。无线调制/解调部分 13 无线调制来自基带调制/解调部分 12 的信号,并将信号输出到无线电天线 11。无线调制/解调部分 13 还对无线电天线 15 11 所接收的信号进行无线解调,并将信号输出到基带调制/解调部分 12。

一个麦克风 14 接收用户的声音并向 AD/DA 转换器 15 输出一个模拟声音信号。AD/DA 转换器 15 将来自音频编解码器 16 的数字信号转换成可被扬声器 17 使用的模拟声音信号,还将来自麦克风 14 的模拟声音信号转换成可被音频编解码器 16 使用的数字信号。音频编解码器 16 压缩或解扩声音信号。扬声器 20 17 再生 AD/DA 转换器 15 所发送的模拟信号。控制器 30 对整个便携式 TV 电话设备进行控制,包括呼叫/呼叫接收控制以及 CG 处理控制。

接下来参考附图 2 和 3 对具有上述设备的便携式 TV 电话设备 (以下被称作 TV 电话) 的呼叫/呼叫接收操作进行描述。

25 如图 2 所示,TV 电话的用户选择是图像传感部件 10 所拍摄的真实影像应该被传送到另一方的终端,还是模仿用户面部的代用影像 (静止图像或运动图像) 应该被传送到另一方的终端 (步骤 S101)。如果步骤 S101 中在预定时间内没有进行选择,则预设的图像被传送到另一方的终端。

选择之后,控制器 30 执行呼叫另一方终端的产生操作 (步骤 S102)。通过中断另一方终端的操作 (响应),语音通信启动,并且一个繁忙状态被设置

(步骤 S103 和 S104)。当语音通信被结束时, 控制器 30 执行从另一方终端断开信道的挂机操作(步骤 S105)。在语音通信期间, 待传送的视频可以被任意转换。

5 如图 3 所示, 一旦被另一方的终端所呼叫(步骤 S201), TV 电话的用户就选择是图像传感部件 10 所拍摄的真实影像应该被传送到另一方的终端, 还是模仿用户面部的代用影像应该被传送到另一方的终端(步骤 S202)。如果在步骤 S202 中预定时间内没有进行选择, 则预设的图像被传送到另一方的终端。

10 选择之后, 控制器 30 执行响应来自另一方终端的呼叫的中断操作(步骤 S203)。当语音通信开始时, 一个繁忙状态被设置(步骤 S204)。当语音通信结束时, 控制器 30 执行从另一方终端断开信道的挂机操作(步骤 S205)。

15 接下来参考附图 4 和 5 对步骤 S104 或 S204 中语音通信期间的图像传送操作进行描述。在语音处理系统中, 如图 4 所示, 基于从麦克风 14 接收到模拟语音信号(步骤 S301), AD/DA 转换器 15 将所接收的模拟声音信号转换成数字声音信号(步骤 S302)。音频编解码器 16 使用一个为 TV 电话所制定的方案(例如 ITU-T G.723.1, 3GPP AMR 等等)来压缩接收自 AD/DA 转换器 15 的数字声音信号。

20 另一方面, 在图象处理系统中, 通过从图像传感部件 10 接收模拟视频信号(步骤 S304), AD/DA 转换器 6 将模拟视频信号转换成数字视频信号, 并且格式转换器 7 将数字视频信号转换成具有 YUV 格式的数字信号, 该格式中包括一个亮度成分 Y 以及色差成分 U 和 V(步骤 S305)。视频编解码器 8 使用为 TV 电话设备制定的方案(例如 ITU-T H.263, ISO/IEC, MPEG4 等等)来对接收自格式转换器 7 的 YUV 信号进行压缩。

25 上述过程在声音和视频处理系统中是同时进行的。多路复用/解复用部分 9 对接收自音频编解码器 16 的声音信号和接收自视频编解码器 8 的视频信号进行多路复用(步骤 S307)。基带调制/解调部分 12 对接收自多路复用/解复用部分 9 的多路复用信号进行基带调制(步骤 S308)。无线调制/解调部分 13 对接收自基带调制/解调部分 12 的信号进行无线调制, 并将信号输出到无线电天线 11(步骤 S309)。无线调制/解调部分 13 所得到的发射电波被从无线电

天线 11 发送到一个基站（步骤 S310）。

以下参考附图 5，对发送由静止图像或运动图像所形成的代用影像来代替真实影像的操作进行描述。图 5 中的所述步骤 S401 到 S403 以及 S407 到 S410 的处理与图 4 中所示步骤 S301 到 S303 以及 S307 到 S310 的相同，对其所做的描述将被省略。

在视频处理系统中，取代图像传感部件 10 拍摄到的真实影像，具有预设 GIF（可交换的图像文件格式）或 JPEG（联合图像编码专家组）格式的数据被从存储在 ROM1 的图形数据（以下称为 CG 数据）中选出并被使用。

限定代用影像的代用影像定义数据具有一种格式，该格式中包括指示有关 CG 数据文件存储位置的指针，该 CG 数据包括“脸部轮廓”，“发型（头发）”，“眉毛”，“眼睛”，“鼻子”，“嘴巴”，“口”，“耳朵”以及“附属物”，例如缎带，部分表示代用影像中“轮廓”、“头发”、“眉毛”、“眼睛”、“鼻子”、“嘴”、“耳朵”以及“附属品”布局位置的信息，以及一个文件长度，如图 6 所示。

如图 7 所示，CG 处理部分 3 在显示设备 5 上向用户提供一个生成窗口来产生代用影像。图 7 所示的生成窗口主要是由一个工具条 18，主绘图窗口 19，部分选择窗口 20 以及选择指针 21 构成的。

用户可以通过用户界面处理部分 4 来对选择指针 21 进行操作。更为特殊的是，当使用选择指针 21 选择工具条 18 上的功能“文件”、“编辑”或“帮助”时，一个下拉菜单被显示，如图 8 所示。当使用选择指针 21 选择子菜单上的功能时，所需的处理可以被执行。当“文件”被选择时，可以对“创建新的代用影像”、“保存文件”、“更新文件”等功能进行选择。当“编辑”被选择时，可以对“撤销”、“复制”、“粘贴”等功能进行选择。当“帮助”被选择时，可以对“检索帮助文件”等功能进行选择。

绘制中的图形被显示在主绘图窗口 19 上。图 7 显示了一种状态，其中脸部轮廓被显示在主绘图窗口 19 中。部分选择窗口 20 是为每个部分，例如“轮廓”、“头发”、“眉毛”、“眼睛”、“鼻子”、“嘴”、“耳朵”以及“其他（附属品）”所准备的。用户使用选择指针 21 从左端的工具箱 22 中选择想要的部分，由此切换部分选择窗口。图 7 显示了一种状态，其中“眼睛”被从这



些部分中选出。

ROM1 中预先为“轮廓”、“头发”、“眉毛”、“眼睛”、“鼻子”、“嘴”、“耳朵”以及“其他(附属品)”中的每一个都准备了多个标准的 CG 数据。这些标准的 CG 数据可以被预先存储在另一个存储设备(第二存储装置)中,而不是 ROM1 中。当给定部分的部分选择窗口 20 被选择时, CG 处理部分 3 显示多个图形,它们是预先为部分选择窗口 20 的显示区域 23 中的被选部分所准备的。用户可以使用选择指针 21 来从被显示的图形中选择一个想要的图形。

如果图形的数量很大,那么其中一些会被显示在显示区域 23 中。当用户使用选择指针 21 移动右端的滚动条时, CG 处理部分 3 根据该移动来滚动显示区域 23。这样,用户就可以看到为选择区域准备的所有图形。

代表光的三种基色 R(红)、G(绿)、B(蓝)灰度级的三个游标 25 被排列在部分选择窗口 20 的下边。用户从显示区域 23 中所显示的那些图形中选择一个想要的,然后使用选择指针 21 移动游标,以选择各种颜色,由此改变被选图形的 RGB 灰度级。通过此操作,可以确定被选图形的颜色。举例来说,当  $(R, G, B) = (0, 0, 0)$  时,颜色是黑色。当  $(R, G, B) = (255, 255, 255)$  时,颜色是白色。

用户可以通过使用选择指针 21 将点击(Grapping)主绘图窗口 19 来对其中的被选图形进行布局。对主绘图窗口 19 来说,一个 X 轴被设置在水平方向,一个 Y 轴被设置在垂直方向。原点  $(X, Y) = (0, 0)$  被设置在左上角。通过这些设置, CG 处理部分 3 可以识别布置有主绘图窗口 19 上图形的坐标。该坐标被用于表示图 6 所示代用影像定义数据中的图形的排列位置。

通过上述过程,为“轮廓”、“头发”、“眉毛”、“眼睛”、“鼻子”、“嘴”、“耳朵”以及“附属品”选择了图形,并确定了颜色。各部分图形被显示在主绘图窗口 19 上,由此模仿用户脸部以生成代用影像。当用户从工具条 18 上的“文件”子菜单中选择“保存”时, CG 处理部分 3 将所生成代用影像各部分的 CG 数据存储在 ROM1 中。CG 处理部分 3 还产生图 6 所示的代用影像定义数据,并将其存储在 ROM1 中,该代用影像定义数据是由指向各部分的 CG 数据的指针,各部分 CG 数据的布局位置,以及文件长度形成的。

在图5的步骤S404中,CG处理部分3从ROM1中提取预先生成并记录的代用影像定义数据。CG处理部分3基于代用影像定义数据中的指针而从ROM1中提取各部分的CG数据,并基于代用影像定义数据中对应的布局位置信息来排列所提取代用影像中的CG数据。各部分的CG数据被合成以产生代用影像数据。CG处理部分3将生成的代用影像数据传送到格式转换器7,以便于将数据转换成具有YUV格式的数字信号(步骤S405)。转换之后,格式转换器7向CG处理部分输出YUV信号。

在真实影像的情况下,视频编解码器8使用为TV电话设备制定的方案(例如ITU-T H.263, ISO/IEC, MPEG4等等)来对CG处理部分3输出的YUV信号进行压缩(步骤S406)。通过步骤S407到S410的处理,压缩信号被发送到另一方的终端。

为了传送运动图像的代用影像而不是静止图像,预先产生并记录了多个代用影像定义数据。CG处理部分3连续为每个基于代用影像定义数据的代用影像定义数据合成CG数据。通过该过程,运动图像的代用影像可以被发送。

也可能不仅传送模仿用户面部的代用影像,同时还传送字符串,例如“动画正在被传送”的CG数据。在这种情况下,字符串的CG数据预先被准备。当指向CG数据和CG数据布局位置的指针被在代用影像定义数据中限时,字符串被添加到代用影像中并被传送。如图9B所示,接收端用户的真实视频被显示在发送端的TV电话上。在接收端的TV电话上,如图9A所示,一个注释26“动画正被传送”被添加到代用影像中,并与发送端用户运动图像的代用影像一起被显示。

为了改变语音通信期间将要发送的代用影像,用户通过用户界面处理部分4对将要发送的视频进行恰当的选择和改变。举例来说,如图10所示,当用户想要发送微笑的面部表示时,他选择按钮“微笑”27。要发送生气的面部表示,用户选择按钮“生气”28。

微笑面部表情的CG数据和代用影像定义数据,生气面部表示的CG数据和代用影像定义数据以及悲伤面部表情的CG数据和代用影像定义数据等等都是预先在ROM1被准备好的。CG处理部分3基于对应于用户选择的代用影像定义数据而对CG数据进行合成,以产生一个代用影像。通过此过程,微笑

面部表情的静止图像或运动图像, 生气面部表情的静止图像或运动图像以及悲伤面部表情的静止图像或运动图像等等可以被发送。

当用户已确定在语音通信期间要发送图像传感部件 10 所拍摄的真实影像时, 他选择图 10 中的按钮“摄影” 29。当对按钮 29 进行操作时, CG 处理部分 3 停止向视频编解码器 8 输出代用影像 (YUV 信号)。同时, 格式转换器 7 开始对将 AD/DA 转换器 6 的输出转换成 YUV 格式的处理。视频处理从图 5 中步骤 S404 到 S406 的处理切换到图 4 中步骤 S304 到 S306 的处理。

接下来参考图 11 对 TV 电话的接收操作进行描述。发自基站并被无线电天线 11 所接收的接收电波是通过无线调制/解调部分 13 而被无线解调的 (步骤 S501 和 S502)。基带调制/解调部分 12 对无线调制/解调部分 13 输出的信号进行基带解调 (步骤 S503)。多路复用/解复用部分 9 将基带调制/解调部分 12 所输出的多路复用信号解复用成一个声音信号和视频信号 (步骤 S504)。

在语音处理系统中, 音频编解码器 16 解扩多路复用/解复用部分 9 所输出的数字声音信号 (步骤 S505)。AD/DA 转换器 15 将解扩的数字声音信号转换成模拟声音信号, 并将该信号输出到扬声器 17 (步骤 S506)。通过这些过程, 发送端的声音被扬声器 17 输出 (步骤 S507)。

另一方面, 在视频处理系统中, 视频编解码器 8 解扩多路复用/解复用部分 9 所输出的数字视频信号 (步骤 S508)。格式转换器 7 对视频编解码器 8 所输出的 YUV 信号的格式进行转换 (步骤 S509)。AD/DA 转换器 6 将接收自格式转换器 7 的数字视频信号转换成模拟视频信号, 并将该信号输出到显示设备 5。通过这些处理, 发送端的视频被显示在显示设备 5 的屏幕上 (步骤 S510)。

在上面的实施例中已对便携式 TV 电话设备举例进行了说明。然而, 本发明并不限制于此。当基带调制/解调部分 12、无线调制/解调部分 13 以及无线电天线 11 被改变成用于任意传输方案的设备时, 那种传输方案的 TV 电话可以被实施。举例来说, 本发明还可以被提供给使用 PSTN (公共开关电话网络) 或 ISDN (综合业务数字网) 的电缆 TV 电话。

另外, 在上面的实施例中, 格式转换部分 7 被用于真实影像数据和代用影像数据的格式转换。可代替的是, 专门用于代用影像数据的格式转换部分可以被安排在 CG 处理部分 3 之中。在这种情况下, CG 处理部分 3 和格式转换器

7 之间代用影像数据的交换并不是必要的。这两个格式转换器受到控制，以便于进行选择性的操作。

如上面所讨论的那样，根据本发明，静止图像，例如肖像的或是运动图像的代用影像，都被发送到语音通信中的一方，而不是讲话者的真实影像，并由此通知语音通信的另一方讲话者很难将其自身显示在屏幕上。因此，讲话者的表情可以通过图像而被传送，同时保护了隐私。另外，由于模仿讲话者面部表示的运动图像被传送，因此具有 TV 电话设备的通信可以被执行，而不会使语音通信的参与者有错误的感觉。

代用影像被转换成 TV 电话设备的视频信号格式并被传送。这样，代用影像可以被发送到另一方的终端，而不需要有另一方终端上专用的再生功能。

存储装置中存储的代用影像的一种和图像传感装置所拍摄的真实视频可以被选择性传送。另外，当字符串与代用影像一起传送时，代用影像和字符串可以同时被显示在通信参与者一边。

被传送的代用影像可以从预先准备的多个代用影像中被选择并被发送。举例来说，可以发送微笑面部表情的代用影像、生气面部表情的代用影像，或是选择性发送悲伤面部表情的代用影像。

作为代用影像成分的一个部分视频被从多个预先准备的部分视频中自由选出，由此产生一个代用影像。因此，代用影像可以通过电话设备自身而被编辑和生成，并被发送。

可以产生一个代用影像，同时随意改变各个部分视频的颜色。更进一步，代用影像中每个部分视频的位置都可以自由设置。

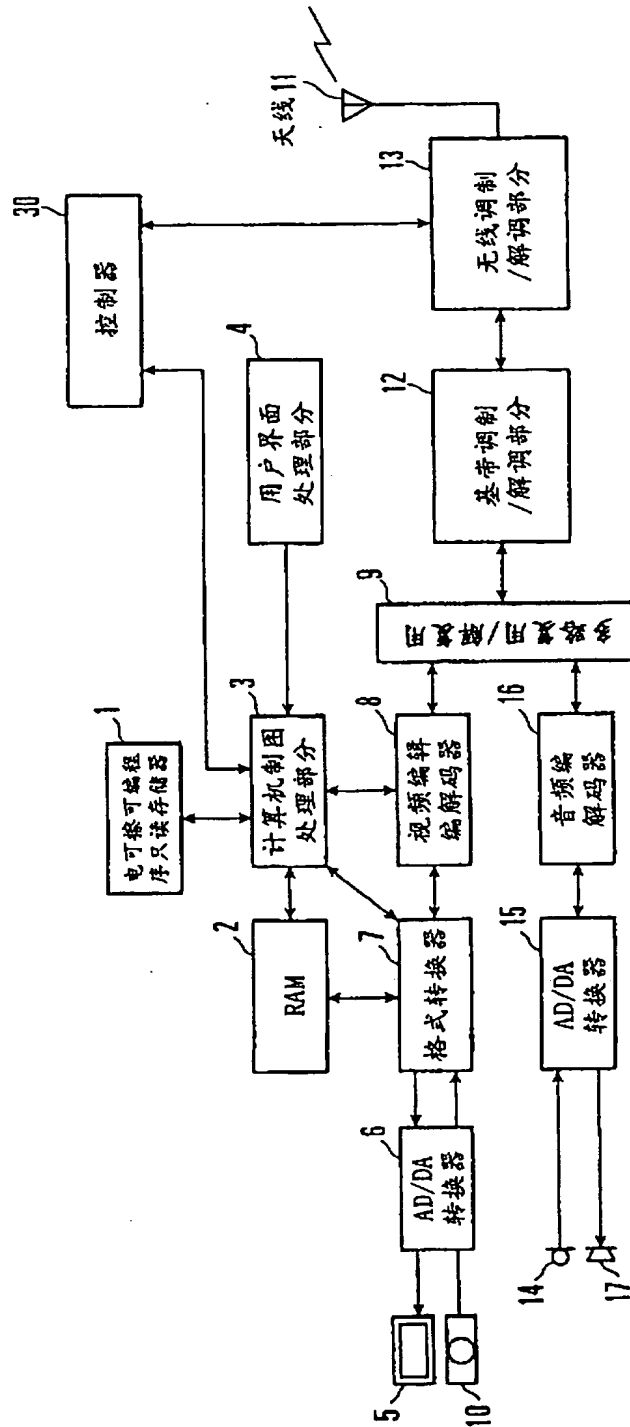


图 1

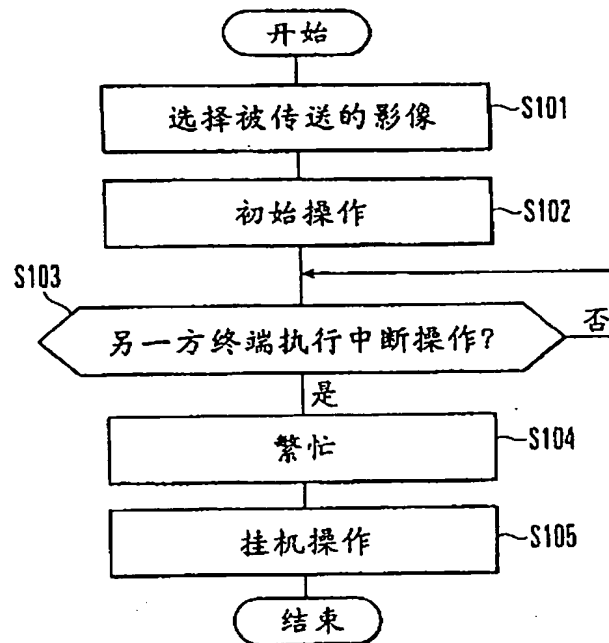


图 2

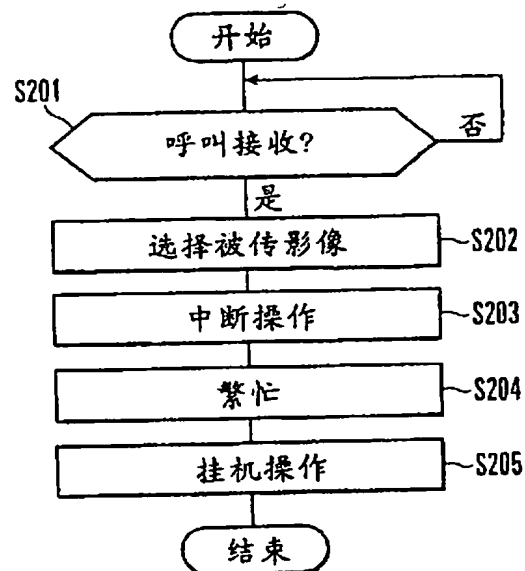


图 3

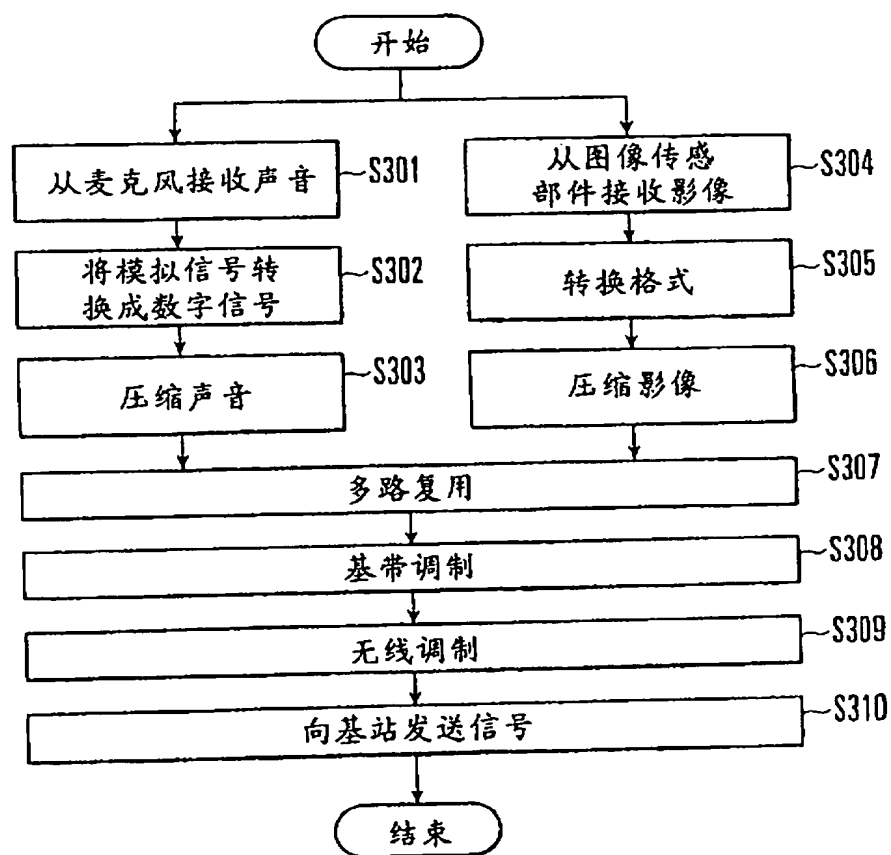
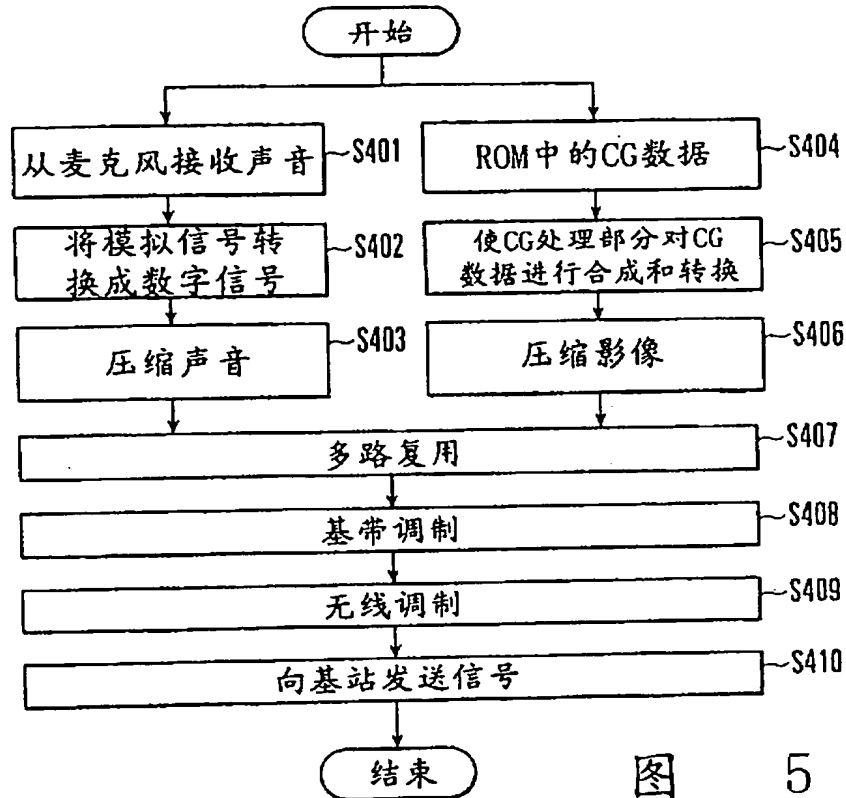


图 4



文件长度	
轮廓文件指针	布局位置
头发文件指针	布局位置
眉毛文件指针	布局位置
眼睛文件指针	布局位置
鼻子文件指针	布局位置
嘴文件指针	布局位置
耳朵文件指针	布局位置
附属品文件#1指针	布局位置
⋮	
附属品文件#N指针	布局位置

图 6



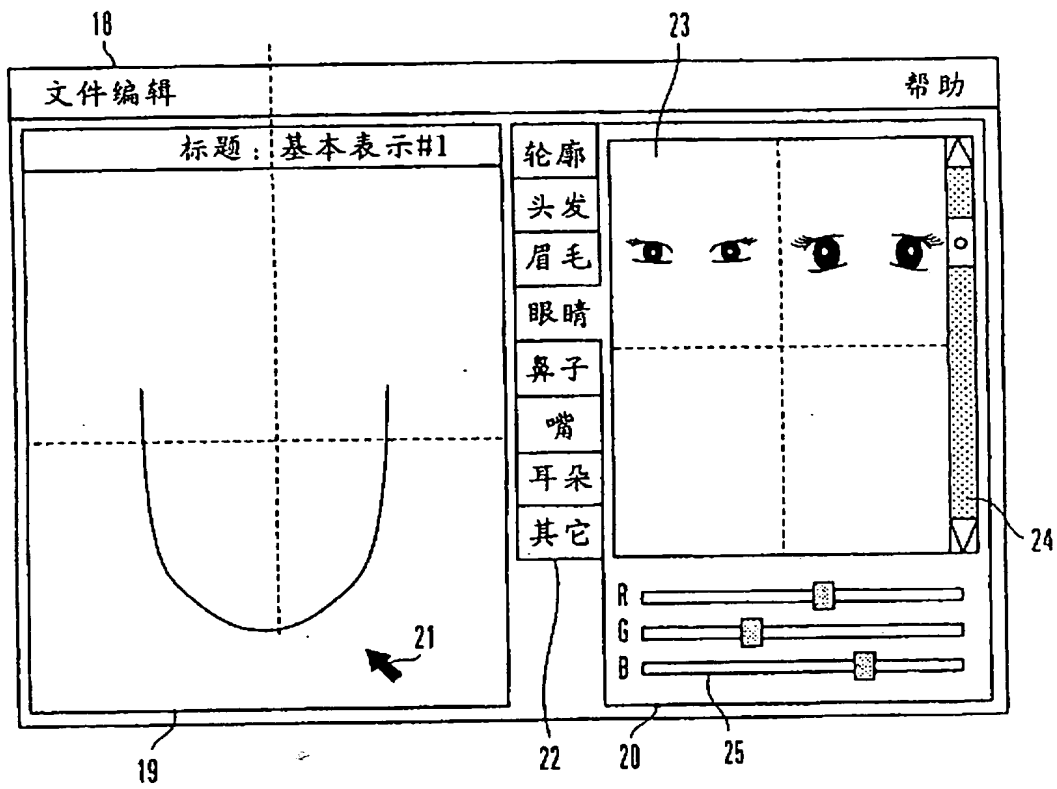


图 7

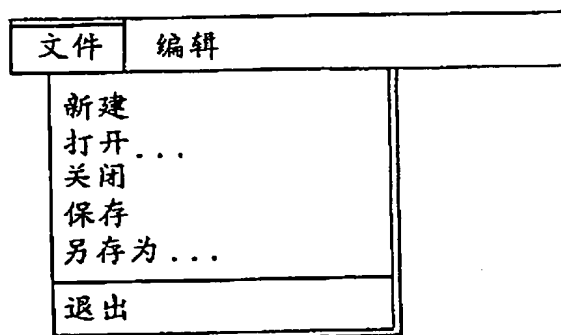


图 8

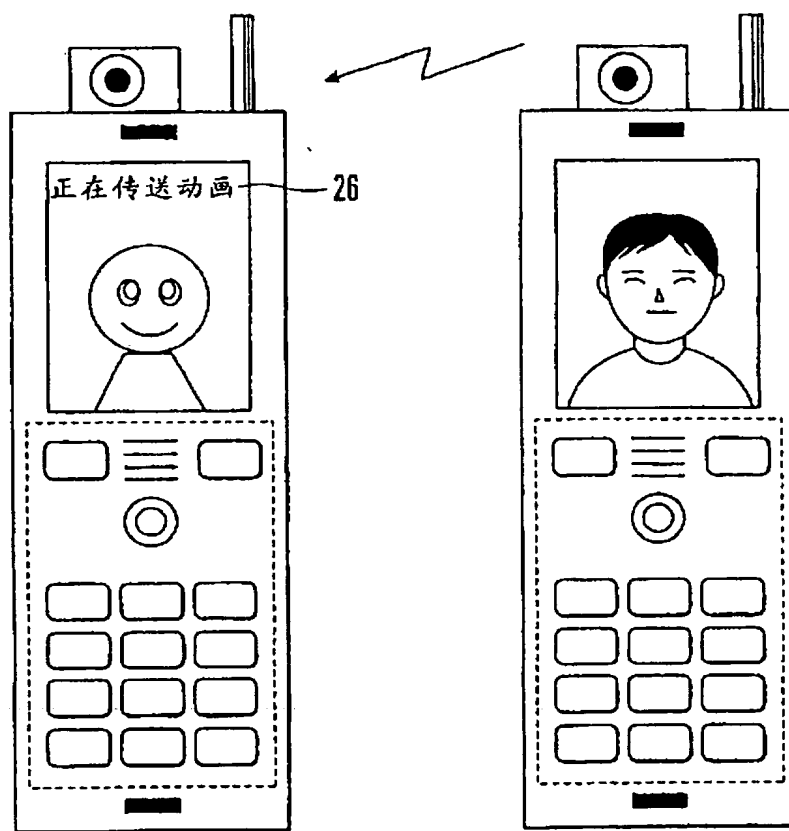


图 9A

图 9B

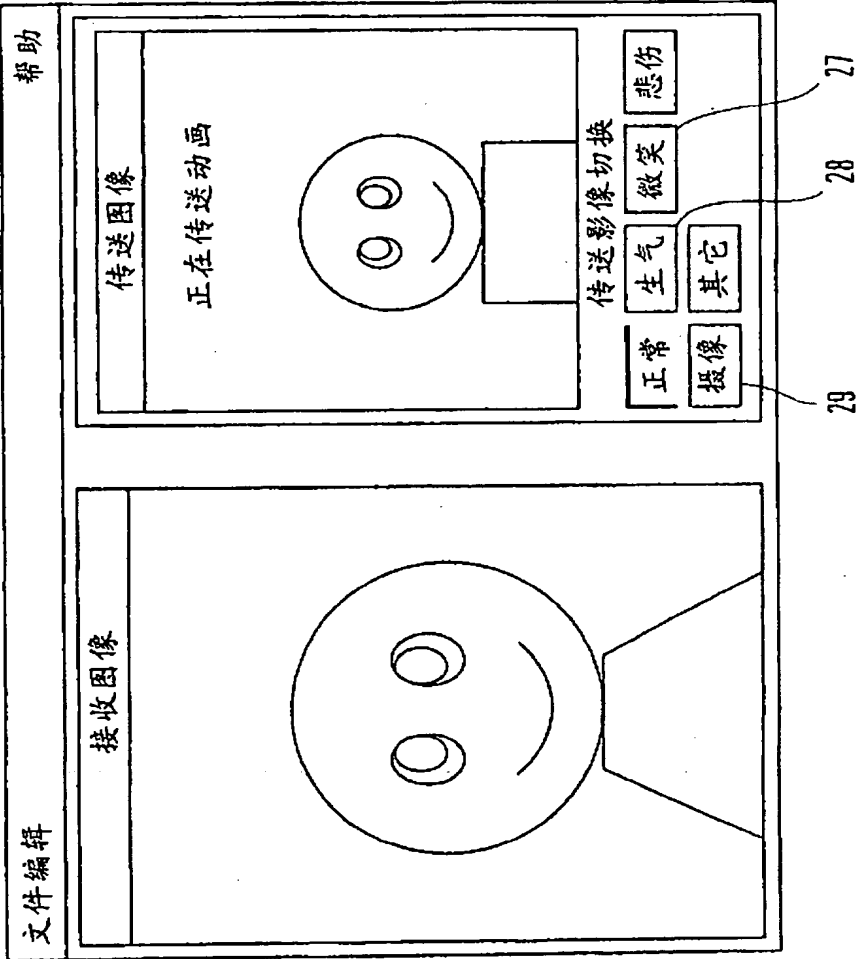


图 10

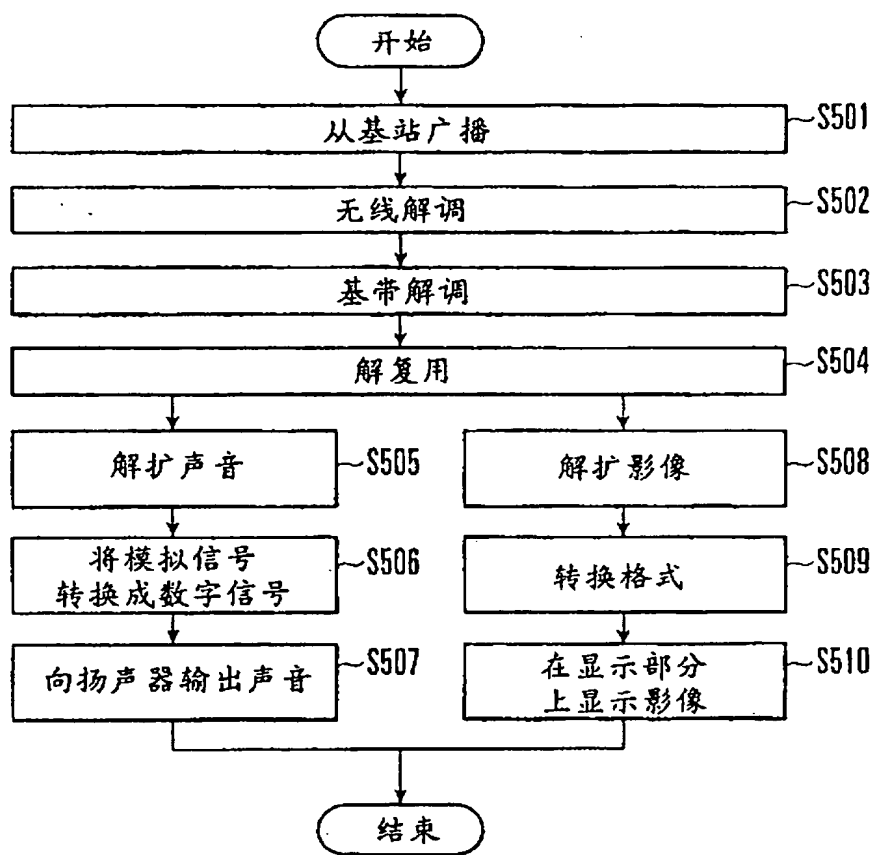


图 11